



Estimation and accounting for the modeling error in probabilistic linearized AVO inversion

Madsen, Rasmus Bødker; Hansen, Thomas Mejer

Published in:
Geophysics

DOI:
[10.1190/geo2017-0404.1](https://doi.org/10.1190/geo2017-0404.1)

Publication date:
2018

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY-NC-ND](#)

Citation for published version (APA):
Madsen, R. B., & Hansen, T. M. (2018). Estimation and accounting for the modeling error in probabilistic linearized AVO inversion. *Geophysics*, 83(2), N15-N30. <https://doi.org/10.1190/geo2017-0404.1>

Estimation and accounting for the modeling error in probabilistic linearized amplitude variation with offset inversion

Rasmus Bødker Madsen¹ and Thomas Mejer Hansen¹

ABSTRACT

A linearized form of Zoeppritz equations combined with the convolution model is widely used in inversion of amplitude variation with offset (AVO) seismic data. This is shown to introduce a “modeling error,” compared with using the full Zoeppritz equations, whose magnitude depends on the degree of subsurface heterogeneity. Then, we evaluate a methodology for quantifying this modeling error through a probability distribution. First, a sample of the unknown probability density describing the modeling error is generated. Then, we determine how this sample can be described by a correlated Gaussian probability distribution. Finally, we develop how such modeling errors affect the linearized AVO inversion results. If not accounted for (which is most often the case), the modeling errors can introduce significant artifacts in the inversion results, if the signal-to-noise ratio is less than 2, as is the case for most AVO data obtained today. However, if accounted for, such artifacts can be avoided. The methodology can easily be adapted and applied to most linear AVO inversion methods, by allowing the use of the inferred modeling error as a correlated Gaussian noise model.

INTRODUCTION

Amplitude variation with offset (AVO) reflection seismic data, and the corresponding amplitude variation with angle (AVA), can be used to analyze how the reflected energy from a layer boundary depends on the offset between a source and a receiver (or the angle of incidence ϕ at the boundary) and the elastic properties around a specific layer boundary (Castagna and Backus, 1993). AVO analysis has been successfully used to identify possible hydrocarbon reservoirs directly from seismic data (Castagna and Backus, 1993; Castagna et al.,

1998). Inversion of NMO corrected prestack seismic data (AVO data) has also been widely used to infer information about the elastic parameters in a deterministic (e.g., Cooke and Schneider, 1983) and a probabilistic framework (e.g., Buland and Omre, 2003).

AVO data are always associated with uncertainty, which can be classified into “measurement uncertainty” and modeling error. The term modeling error (sometimes also referred to as “modelization error” or “theoretical error”) is, in this paper, referring to the use of an inexact theory in the modeling (prediction) of the result of measurements as opposed to errors arising from inaccurate measurements due to, e.g., instrument errors (Tarantola and Valette, 1982b; Sen and Stoffa, 1996; Downton, 2005; Tarantola, 2005).

An abundance of work exists that acknowledges modeling errors related to AVO modeling (Gerstoft and Mecklenbräuker, 1998; Buland and Omre, 2003; Riedel et al., 2003; Dosso and Holland, 2006; Chen et al., 2007; Rabben et al., 2008; Bosch et al., 2010; Aune et al., 2013).

A widely considered modeling error with respect to AVO data is related to the use of approximations to the Zoeppritz (1919) equations, which allow the AVO/AVA response at a plane-layer interface to be computed analytically. These approximations are typically based on the small-contrast approximation given by Aki and Richards (1980) and are further developed by, e.g., Shuey (1985) and Castagna and Backus (1993). These approximations are used for AVO analysis and inversion.

Most approximations of the Zoeppritz equations are assumed to be valid (implying insignificant modeling errors) for incidence angles $\theta < 30^\circ$ (Shuey, 1985; Castagna and Backus, 1993; Buland and Omre, 2003; Mavko et al., 2009). If the variations in the elastic properties are assumed to be very smooth, implying small contrasts, the approximation may be valid for incidence angles $\theta > 30^\circ$. Furthermore, by using the average angle between the incidence and transmitted angle, the modeling error can be decreased in general (Downton and Ursenbach, 2006).

Even for small incidence angles (less than 30°), systematic errors have been shown to arise in linearized AVO inversion due to the

Manuscript received by the Editor 26 June 2017; revised manuscript received 19 September 2017; published ahead of production 21 November 2017; published online 11 January 2018.

¹University of Copenhagen, Niels Bohr Institute, Copenhagen, Denmark. E-mail: rasmus.madsen@nbi.ku.dk; tmeha@nbi.ku.dk.

© 2018 Society of Exploration Geophysicists. All rights reserved.

modeling error of applying a linear operator as a forward model (Downton, 2005; Rabben et al., 2008). The problem is most severe using two-term approximations (Downton, 2005). In this case, significant systematic errors for the gradient estimate can be detected. Further, linear approximations tend to produce false predictions, or artifacts in noncontinuous and nonsmooth subsurface models (Stolt and Weglein, 1985).

Thus, modeling errors related to using approximations to the Zoeppritz equations have been widely considered. Yet, little work has been done to quantify this modeling error and to account for this modeling error as part of inverting AVO data.

Here, we develop the approach proposed by Hansen et al. (2014) to simulate, model, and account for modeling errors in relation to AVO modeling. Specifically, and as a first example, the widely used small-contrast approximation of the Zoeppritz equations will be investigated.

We expect such modeling errors to represent a lower limit of the full set of modeling errors inherent in AVO data. A variety of sources for modeling errors exist in seismic data, such as using a 1D convolutional model to reflect a 3D physical system, the use of the acoustic-wave equation as opposed to the anisotropic visco-elastic wave equation, imperfections in data processing, general anisotropy considerations, the effects of processing the raw data, the coupling of data to physics within the forward model, uncertain wavelet estimates, and uncertainty on the low-frequency model. (see also Ball et al., 2015; Li et al., 2015; Thore, 2015). For example, we expect higher magnitude modeling errors related to using the Zoeppritz equations as opposed to using the full wave equation. However, as shown in this paper, even disregarding this lower limit of modeling errors can potentially lead to significant biases when such AVO data are inverted without accounting for modeling errors.

First, we introduce the approximate forward model and describe the modeling error associated with it. This is followed by a discussion on how to estimate and quantify this error. This is done in a probabilistic framework. First, a set of realizations representing a sample of a (unknown) probability distribution describing the modeling error is generated. Then, it is demonstrated that this sample can be reasonably described by a multivariate Gaussian probability density. Finally, we investigate the possibility of accounting for (and ignoring) the forward modeling error during probabilistic linear inversion of seismic AVO data (as in Buland and Omre, 2003), considering different levels of measurement uncertainty.

LINEARIZED AVO FORWARD MODELING

In the following, we will be using the term AVO as synonymous with AVA. Perhaps, the most correct approach to solve the AVO forward problem (i.e., simulating an AVA gather) is to use some form of full-waveform modeling to simulate shot gathers, followed by a NMO correction and sorting according to angle of incidence ϕ , to obtain a set of AVA gathers. It is however computationally demanding, and especially if used as part of an inverse problem, it becomes impractical (Virieux and Operto, 2009). It is also argued that in real-world cases, processing the raw seismic data such that it enables the use of a simpler forward problem is often advantageous (Claerbout et al., 2004). AVO data are, for example, processed such that each data point can be associated with the reflection at a specific point in the subsurface. The Zoeppritz equations allow an analytical relation between the amplitude of reflections and elastic

parameters around an interface in the subsurface. In practice, a widely used forward model to simulate AVO data is a combination of a linearized approximation to Zoeppritz equations combined with the convolution model.

The success of the simpler approximations can be attributed to their ease of evaluation and interpretation (Castagna and Backus, 1993; Mavko et al., 2009), and their ability to linearize the AVO inverse problem (Buland and Omre, 2003).

1D linearized AVO forward modeling

The Zoeppritz (1919) equations describe the full set of angle-dependent reflectivities at a plane interface between two media with different elastic properties for a plane wave.

The Aki and Richards forward model

Aki and Richards (1980) propose a small-contrast approximation to Zoeppritz equations, in which the reflection coefficient, as a function of incidence angle ϕ can be computed from the elastic parameters above and below the interface using

$$R(\phi) \approx a_{v_p}(\phi) \frac{\Delta v_p}{\bar{v}_p} + a_{v_s}(\phi) \frac{\Delta v_s}{\bar{v}_s} + a_\rho(\phi) \frac{\Delta \rho}{\bar{\rho}}, \quad (1)$$

where the coefficients are given by

$$a_{v_p}(\phi) = \frac{1}{2 \cos^2 \phi}, \quad (2)$$

$$a_{v_s}(\phi) = -\frac{4\bar{v}_s^2}{\bar{v}_p^2} \sin^2 \phi, \quad (3)$$

$$a_\rho(\phi) = \frac{1}{2} \left(1 - 4 \frac{\bar{v}_s^2}{\bar{v}_p^2} \sin^2 \phi \right), \quad (4)$$

where \bar{v}_p , \bar{v}_s , and $\bar{\rho}$ represent the average P-wave, S-wave, and density over the interface, whereas Δv_p , Δv_s , and $\Delta \rho$ represent elastic contrasts over the interface. Equation 1 is valid for a small percentile change in elastic properties, i.e., for a small contrast between layers.

The assumption of small contrasts is linked to the relative change in contrasts $\Delta v_p/\bar{v}_p$, $\Delta v_s/\bar{v}_s$, and $\Delta \rho/\bar{\rho}$. We will refer to this choice of AVO forward model of reflectivities as the “Aki and Richards forward model.”

Note that equation 1 is a better approximation if the average angle at the interface is used, as opposed to the incidence angle (Downton and Ursenbach, 2006).

The Buland and Omre forward model

Buland and Omre (2003) adapt the following forward approximation proposed by Stolt and Weglein (1985), which expands the Aki and Richards approximation, in which reflection coefficients are now also time dependent:

$$R(t, \phi) \approx a_{v_p}(t, \phi) \frac{\partial}{\partial t} \ln v_p(t) + a_{v_s}(t, \phi) \frac{\partial}{\partial t} \ln v_s(t) + a_\rho(t, \phi) \frac{\partial}{\partial t} \ln \rho. \quad (5)$$

The difference terms in equation 1 are in equation 5 substituted with the partial derivative of the logarithmic value of each material parameter, e.g., $\partial/\partial t \ln v_p(t)$ replaces $\Delta v_p/\bar{v}_p$. This substitution is valid only for small contrasts in the elastic parameters.

The time dependency in the coefficients $a_{v_p}(t, \phi)$, $a_{v_s}(t, \phi)$, and $a_\rho(t, \phi)$ is a generalization of the Aki and Richards coefficients (equations 2–4) with the time-dependent averages \bar{v}_p , \bar{v}_s , and $\bar{\rho}$. To solve this equation, it is here assumed that these averages are described by a known background model.

This allows a linear relation between the derivative of the model parameters (the logarithm of the elastic parameters) and the reflection coefficients

$$\mathbf{R} = \mathbf{A}\mathbf{D}\mathbf{m}, \quad (6)$$

where \mathbf{A} is the linear operator composed of the coefficients $a_{v_p}(t, \phi)$, $a_{v_s}(t, \phi)$, and $a_\rho(t, \phi)$ from equation 5 and \mathbf{D} is the derivative matrix. We will refer to this choice of AVO forward model for reflectivities as the “Buland and Omre forward model.”

The linear convolution model

A seismic trace is obtained using the convolution model:

$$S(t) = W(t) * R(t) \equiv \int_0^{t_s} W(\tau) R(t - \tau) d\tau, \quad (7)$$

where $S(t)$ is the seismic trace, $R(t)$ is the reflectivity series (earth response), such as given in for example equations 1 and 5, and $W(t)$ is the wavelet (source-time function), where t_s is the duration of the source input. Evaluating the AVO forward problem using the Zoeppritz equations (to obtain a seismic trace $S_{\text{zoepp}}(t)$) thus amounts to first computing the reflection coefficients using the Zoeppritz

equations, followed by a convolution with a wavelet. This is a non-linear process.

By applying the convolution model for several angles of incidence, each with a possibly unique wavelet, a linear relationship between AVO seismic data and the elastic model parameters using equation 6 is written as

$$\mathbf{d}_{\text{AVO}} = \mathbf{W}\mathbf{R} = \mathbf{W}\mathbf{A}\mathbf{D}\mathbf{m}, \quad (8)$$

where \mathbf{W} is a convolution matrix containing the wavelet. Buland and Omre (2003) introduce a Bayesian linearized AVO inversion technique by adapting the linear relation in equation 8 when the wavelet is known. This relation is considered the “full Buland and Omre AVO forward model” and allows a full description between model parameters and AVA data using a linear theory.

Calculating the forward modeling error

For one specific 1D elastic model, the modeling error S_{error} (related to a specific elastic model) is calculated as the difference (residual) between the seismic signal from the full Zoeppritz equations $S_{\text{zoepp}}(t)$ and the seismic signal from the approximate forward model $S_{\text{app}}(t)$:

$$S_{\text{error}}(t) = S_{\text{zoepp}}(t) - S_{\text{app}}(t), \quad (9)$$

$$= W(t) * R_{\text{zoepp}}(t) - W(t) * R_{\text{app}}(t), \quad (10)$$

where $R_{\text{zoepp}}(t)$ is the reflectivity series calculated with the Zoeppritz equations and $R_{\text{app}}(t)$ is the one calculated with an approximate forward model (equation 1 or 6). For a single realization from a Gaussian probability density identical to “well B” in Buland and Omre (2003), the AVO forward response calculated using Zoeppritz equation and the Buland and Omre forward model are shown alongside the modeling error (equation 9) in Figure 1. The modeling error is increasing with angle of incidence.

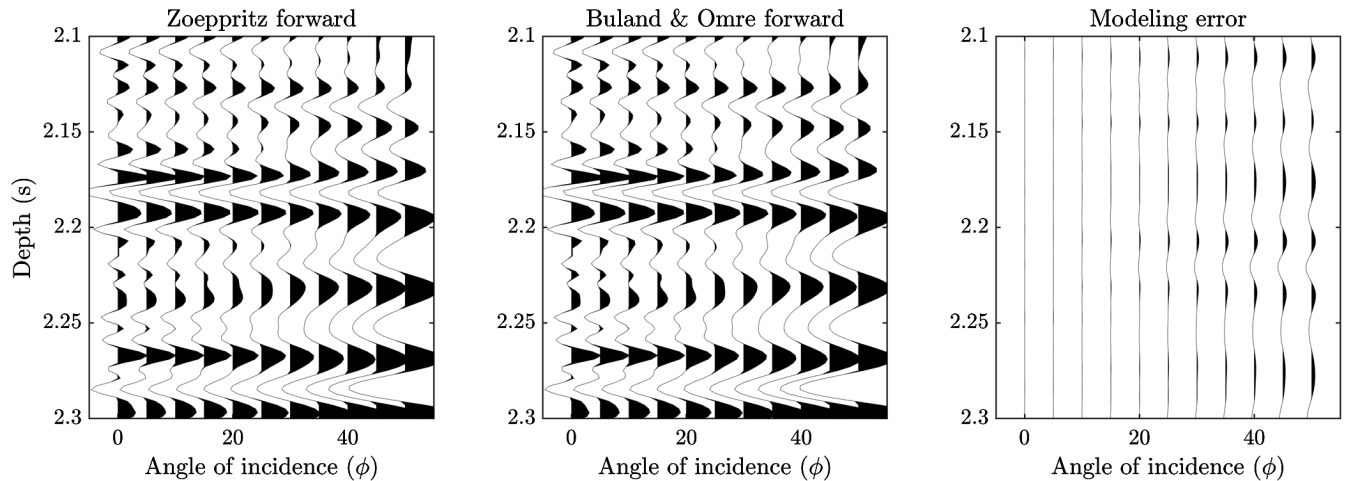


Figure 1. AVO (AVA) data obtained using the full Zoeppritz and the Buland and Omre forward model for a realization from a Gaussian probability density (the small reflectivity model) identical to that of Buland and Omre (2003). On the right, the corresponding forward modeling error is shown. The elastic model used to generate these forward responses is shown in Figure 2.

QUANTIFYING THE FORWARD-MODELING ERROR

Hansen et al. (2014) demonstrate how to quantify the modeling error probabilistically through a probability density, $\theta(\mathbf{d}|\mathbf{m})$. The main idea is to generate a large sample of an assumed (unknown) probability density reflecting the modeling error. To do this, a source of the modeling error has to be identified, and quantified probabilistically, such that realizations of the probability density describing the source of the modeling error are generated. An algorithm that can sample the probability density will suffice, and the actual probability density itself need not be known. The source of the modeling error can for example be subsurface variability and uncertainty related to the wavelet.

In some cases, the obtained sample of the modeling error can be described by a Gaussian probability, in which case, a full description of a Gaussian modeling error can be estimated from the sample of the modeling error as a mean and a covariance (Hansen et al., 2014). If the Gaussian model is adopted, then it can account for the Gaussian modeling error in a linear inverse Gaussian problem, such as the one considered by Buland and Omre (2003), by addition of the mean and the covariance of the measurement uncertainty and the modeling error (for details, see Mosegaard and Tarantola, 1995; Tarantola, 2005).

Forward-modeling error

As an example, we consider two different types of geostatistical subsurface models to illustrate how to sample and quantify the associated modeling error. These two models, a small-contrast and a large-contrast model, respectively, represent extreme cases of subsurface variability. Figure 2 shows a realization from both models. Arbitrarily large amounts of realizations of these geostatistical models are generated that all respect the assumed statistical properties.

The “small-contrast” model is identical to the one presented by Buland and Omre (2003), and it is defined as a Gaussian probability density. The correlation between the elastic parameters is 0.7. The choice of a Gaussian-type covariance model results in smooth models, in which the contrasts in elastic parameters at an interface are relatively small. The small-contrast prior model can therefore be regarded as a “best-case” scenario for the forward approximations based on the small-contrast approximation.

The other extreme is a “large-contrast” prior model, in which contrasts in the elastic parameters are high. The model is based on a truncated plurigaussian prior distribution that allows simulation of complex arrangements of lithofacies (Armstrong et al., 2011). As seen in the lower plot in Figure 2, a set of four unique discrete layers are simulated with rapid shifts in lithofacies. The elastic parameters of the large-contrast prior emulate a simplified model of what is

found in North Sea green-sand environments (Svendsen et al., 2012). The large-contrast prior model may represent a “worst-case” scenario for the use of the small-contrast approximation.

In a real geologic setting, the elastic parameters of the subsurface are expected to be somewhere in between these two extremes.

The convolution model requires the wavelet to be known. Wavelet estimation is an inverse problem in itself. Here, it is assumed that the wavelet is known and is a Ricker wavelet with linearly decreasing center frequency from 50 Hz at zero offset to 25 Hz at the largest incidence angle ($\phi = 50^\circ$). The interval between each trace is $\Delta\phi = 5^\circ$. This yields 11 traces. For 100 time samples, the data (\mathbf{d}_{AVO}) are of size $N_d = 1100$. An example of synthetic data calculated using the small-contrast prior is shown in Figure 1. In total, 100 time samples times the three elastic parameters gives $N_m = 300$ model parameters ($\mathbf{m} = [\ln(v_p)^T, \ln(v_s)^T, \ln(\rho)^T]^T$).

Using one realization from either model, one realization of the modeling error $\mathbf{d}_e = [S_{\text{error},\phi=0}, S_{\text{error},\phi=5}, \dots, S_{\text{error},\phi=50}]$ is calculated using equation 9. One thousand of these modeling error realizations are computed, which represent a sample of the (unknown) probability distribution describing the modeling error. The average standard deviation of this sample at different incidence angles is calculated for the Buland and Omre forward (equation 8) and the Aki and Richards forward (equation 1 plus 7), and it is shown in Figure 3. For the Aki and Richards forward, the average angle ϕ_{avg} between the incidence angle and the transmitted angle is used instead of the incidence angle, as described by

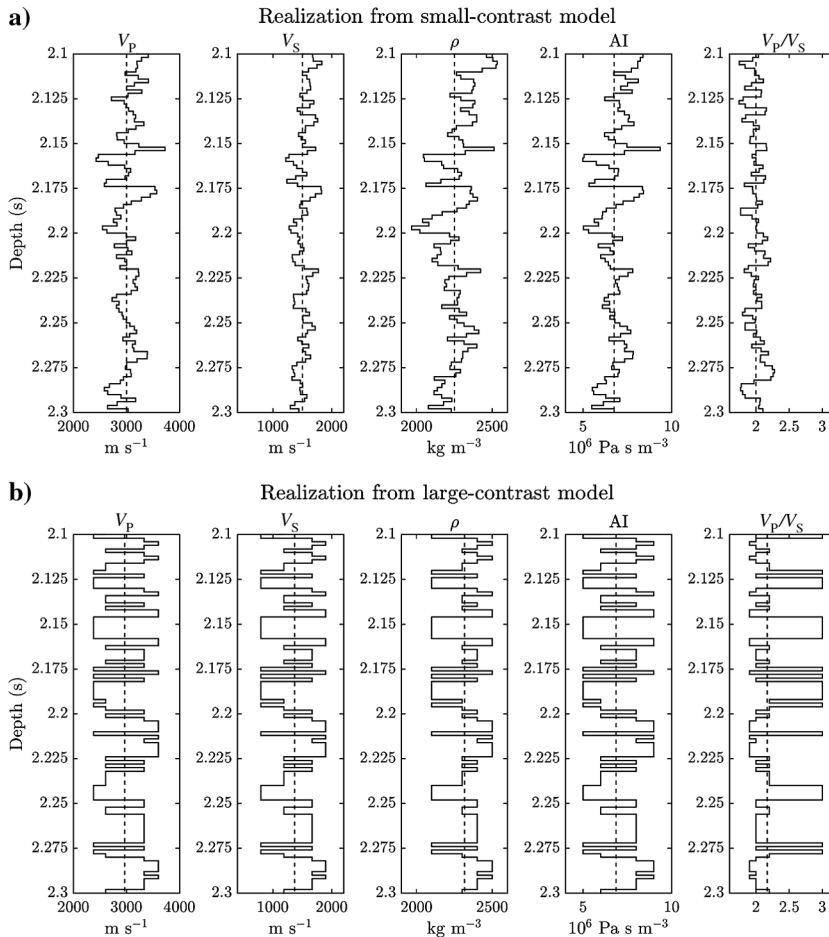


Figure 2. (a) One realization from the small-contrast statistical model and (b) the large-contrast statistical model considered.

Downton and Ursenbach (2006). For the Buland and Omre forward, results are shown using the incidence angle and the average angle.

The average standard deviation introduces a way to measure the amplitude of the modeling error at different incidence angles. For a relative measure on the severity of the modeling error, we compare with the magnitude of the data signal. The data signal is obtained as the average standard deviation of a sample of 1000 forward realizations using the Zoeppritz equations (i.e., to obtain the average standard deviation of the signal, not the error). Comparing this signal (black line) with the modeling error of each approximation (dashed, dotted, and light-gray line) in Figure 3, a visual signal-to-noise ratio (S/N) is obtained.

Both forward approximations show increasing amplitude of the modeling error with the increasing angle of incidence, in agreement with Figure 1. For the small-contrast model (the uppermost plot), the modeling error of the Buland and Omre forward model (the dotted and dashed lines) is higher than the Aki and Richards forward model (the light gray line) for all incidence angles. This is due to the necessary additional small-contrast assumption (logarithmic approximation) in the Buland and Omre forward model. Especially, for angles of incidence $\phi > 35^\circ$, the S/N becomes low for the Buland and Omre forward model because the average modeling error rises. Meanwhile, the average S/N is relatively high for the Aki and Richards forward model due to a generally low modeling error.

Similar trends are visible for the large-contrast model (the lowermost plot). However, the S/N is considerably lower and more significant for both forward models. This is because both forward models rely on the assumption of small contrasts in the elastic parameters, which is not provided by the large-contrast prior model. The S/N is still high for both models for lower angles of incidence ($\phi < 20^\circ$). At larger angles of incidence, the S/N becomes considerably lower for both forward models. This culminates in an S/N = 1 for the Buland and Omre forward model with incidence angle (dotted line) for angles of incidence $\phi > 40^\circ$. This implies that the average modeling error has the same amplitude as the seismic signal for these wide incidence angles. Interestingly, the average modeling error is actually higher for the forward using an average angle (the dashed and light gray lines) for angles of incidence between $10^\circ < \phi < 35^\circ$. A possible explanation for this curiosity is that a phase component arises in the solution when using the average angle in the forward models for large contrasts in the elastic parameters. To remedy this, we use the magnitude of the reflection coefficient's real and complex part as suggested by Lay and Wallace (1995).

In general, the most erroneous of the two approximations is the Buland and Omre forward model when using the incidence angle (the dotted line). However, because the Buland and Omre forward allows a linear relationship between AVO seismic data and the elastic model parameters (equation 8), consequently allowing a Bayesian

linearized AVO inversion; we will be using this forward model in the following calculations.

The results also indicate that using the average angle, instead of the incidence angle, reduces the modeling error substantially, as suggested by Downton and Ursenbach (2006). But, as will be discussed later, the average angle is typically not known when performing linearized AVO inversion because the elastic parameters that are needed to compute the transmission angle are not known prior to inversion. Further, if an elastic a priori model is assumed, it is most often assumed to be smoothly varying, such that there will be little difference between the angle of incidence and the average angle.

A Gaussian model of the forward-modeling error

Assuming that the modeling error is Gaussian, a Gaussian model in the form of a mean and a covariance $\mathcal{N}(\mathbf{d}_{Tapp}, \mathbf{C}_{Tapp})$ describing the forward-modeling error $\theta(\mathbf{d}|\mathbf{m}) \sim \mathcal{N}(\mathbf{d}_{Tapp}, \mathbf{C}_{Tapp})$ can be constructed from a large sample of modeling error realizations, as proposed by Hansen et al. (2014) (see Appendix A). The estimated covariance matrix \mathbf{C}_{Tapp} of the modeling error using the Buland and Omre forward approximation is depicted in Figure 4 for the two prior models. Both covariance matrices show a heavily banded

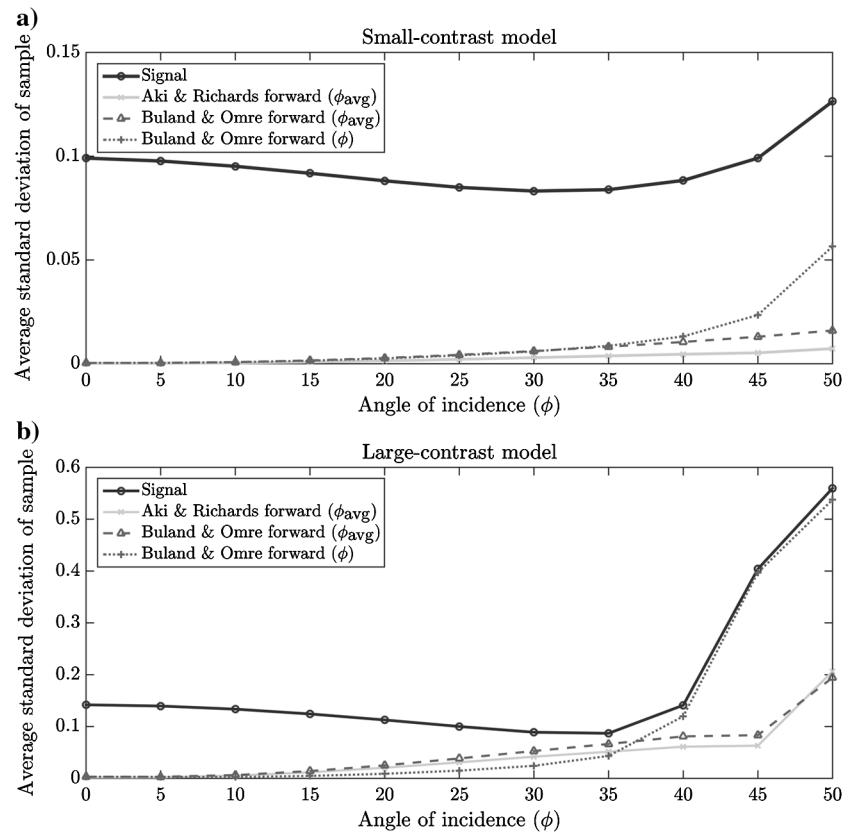


Figure 3. Average standard deviation of the modeling error applying the Aki and Richards approximation using the average angle (the light gray line), and the Buland and Omre approximation using the average and the incidence angle (dashed and dotted line) plotted against the incidence angle. The lines represent the average of 1000 realizations of the modeling error. Also plotted is the forward response from 1000 subsurface realizations using the Zoeppritz equations (the black line) as the signal. Results are shown for (a) the small-contrast model and (b) the large-contrast model.

structure, which indicates a high degree of correlated modeling error. The bands are especially distinguishable at intervals of 100 data points, indicating the correlation between the same time samples at different incidence angles. The covariance matrix also shows that adjacent time samples are correlated within these bands. The amplitude of the modeling error increases along the diagonal, i.e., with an increasing angle of incidence, as is also shown in Figure 3.

In addition to the estimated Gaussian model $\mathcal{N}(\mathbf{d}_{Tapp}, \mathbf{C}_{Tapp})$, we also examine a Gaussian model that contains only the uncorrelated part of the estimated covariance matrix. This is done to assess the importance of the apparent correlation of the modeling error and to evaluate the common practice, as mentioned earlier, of describing the overall data uncertainty with an uncorrelated Gaussian model with a known variance. The full estimated Gaussian model (using equations A-3 and A-4) containing the correlated and uncorrelated parts will henceforth be known as \mathbf{C}_{Tapp1} , and the model containing only the uncorrelated part will be known as \mathbf{C}_{Tapp2} . However, the question now remains: Is the Gaussian model assumption for the modeling error reasonable? This is investigated further in the following by assessing the Gaussian assumption qualitatively and quantitatively.

Qualitative assessment — Visual comparison

A qualitative assessment of the validity of the Gaussian model assumption $\theta(\mathbf{d}|\mathbf{m}) \sim \mathcal{N}(\mathbf{d}_{Tapp}, \mathbf{C}_{Tapp})$ is obtained by visually comparing realizations of the observed (actual) modeling error (obtained using equation 10), with realizations of the Gaussian model describing the modeling error, which should show similar characteristics. Three realizations from \mathbf{C}_{Tapp1} and \mathbf{C}_{Tapp2} are simulated and plotted in Figure 5 alongside the observed modeling error. Realizations are shown for the small-contrast and large-contrast prior model.

Realizations from \mathbf{C}_{Tapp1} generally show the same characteristics as the observed modeling error for both prior cases. Visually, it seems that the amplitudes and patterns are alike for corresponding

incidence angles of the observed and simulated realizations. The main variability of the modeling error seems to be imitated and it is very hard, if not impossible, to distinguish between the observed modeling error and the realizations of \mathbf{C}_{Tapp1} . This suggests qualitatively that the choice of a Gaussian model to represent the modeling error is valid. On the contrary, the realizations from \mathbf{C}_{Tapp2} are not able to reproduce the pattern of the observed modeling error for either prior model. The amplitudes are similar to the observed modeling error, but the white noise is a poor imitation of the overall characteristics.

In summary, a qualitative assessment of the modeling errors through visual comparison shows that the distribution and frequency content of the observed modeling error is fairly well-represented by the estimated $\mathcal{N}(\mathbf{d}_{Tapp1}, \mathbf{C}_{Tapp1})$. A quantitative assessment revealed that the correlated Gaussian model cannot completely describe the observed modeling error, but that it is a much better representation than a simple uncorrelated Gaussian model (see Appendix B). The tails of the observed modeling error $\mathbf{d}_{e,obs}$ in the 1D marginal distribution could potentially be fitted better with a distribution capable of producing outliers more regularly (Cauchy distribution, Voigt profile, etc.). Whether the Gaussian description of the modeling error is “good enough” will ultimately be determined by what it will be used for, as, for example, solving the inverse AVO problem.

FORWARD-MODELING ERROR IN BAYESIAN LINEARIZED AVO INVERSION

The linear relation in equation 8 allows formulating a linear inverse problem, in which the elastic parameters are estimated directly from AVO data. Buland and Omre (2003) solve this inverse problem in a probabilistic framework following Tarantola and Vallette (1982a). Here, the noise on the AVO data is assumed to follow a Gaussian distribution with a mean of zero and a covariance matrix of \mathbf{C}_D ($\epsilon \sim \mathcal{N}(0, \mathbf{C}_D)$). The prior information on the logarithm of

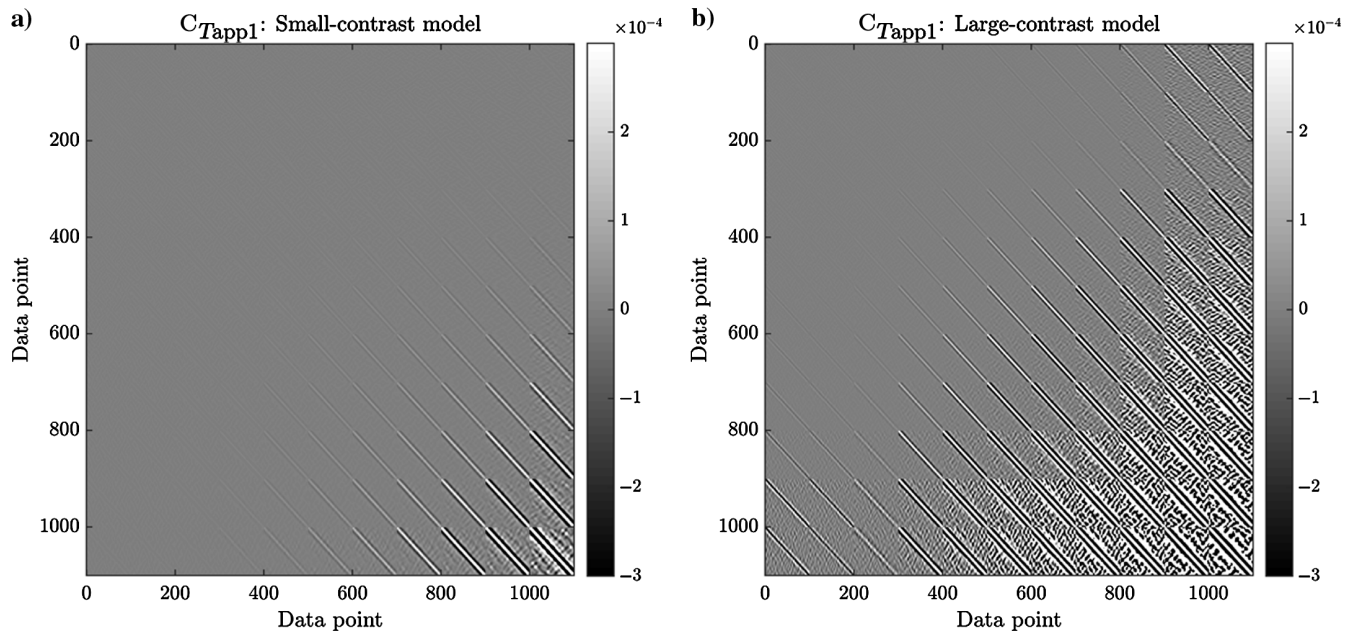


Figure 4. Estimated covariance matrices \mathbf{C}_{Tapp} of the forward-modeling error for both prior models. The covariances are estimated from a large sample (1000 realizations) of the modeling error. The same color scaling is used for both plots. The covariances show a banded structure, indicating that the estimated modeling error is correlated.

the elastic parameters is assumed to follow a Gaussian distribution with mean vector $\boldsymbol{\mu}_M$ and covariance matrix \mathbf{C}_M ($\mathbf{m} \sim \mathcal{N}(\boldsymbol{\mu}_M, \mathbf{C}_M)$), identical to the small-contrast statistical model presented earlier. The posterior probability density of the model parameters $\tilde{\mathbf{m}}$ is described as a Gaussian probability distribution $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$ with mean

$$\tilde{\mathbf{m}} = \boldsymbol{\mu}_M + (\mathbf{WADC}_M)^\top \mathbf{C}_D^{-1} (\mathbf{d}_{\text{obs}} - \mathbf{WAD}\boldsymbol{\mu}_M) \quad (11)$$

and covariance

$$\tilde{\mathbf{C}}_M = \mathbf{C}_M - (\mathbf{WADC}_M)^\top \mathbf{C}_D^{-1} \mathbf{WADC}_M, \quad (12)$$

where \mathbf{d}_{obs} is the observed data.

The least-squares solution described in equations 11 and 12 allows taking the modeling errors into account quite easily, as long as the modeling error can be described by a Gaussian probability density (Mosegaard and Tarantola, 2002; Tarantola, 2005). In that case, the mean (here zero) and covariance describing the measurement and modeling error combine through addition of the mean and covariances as

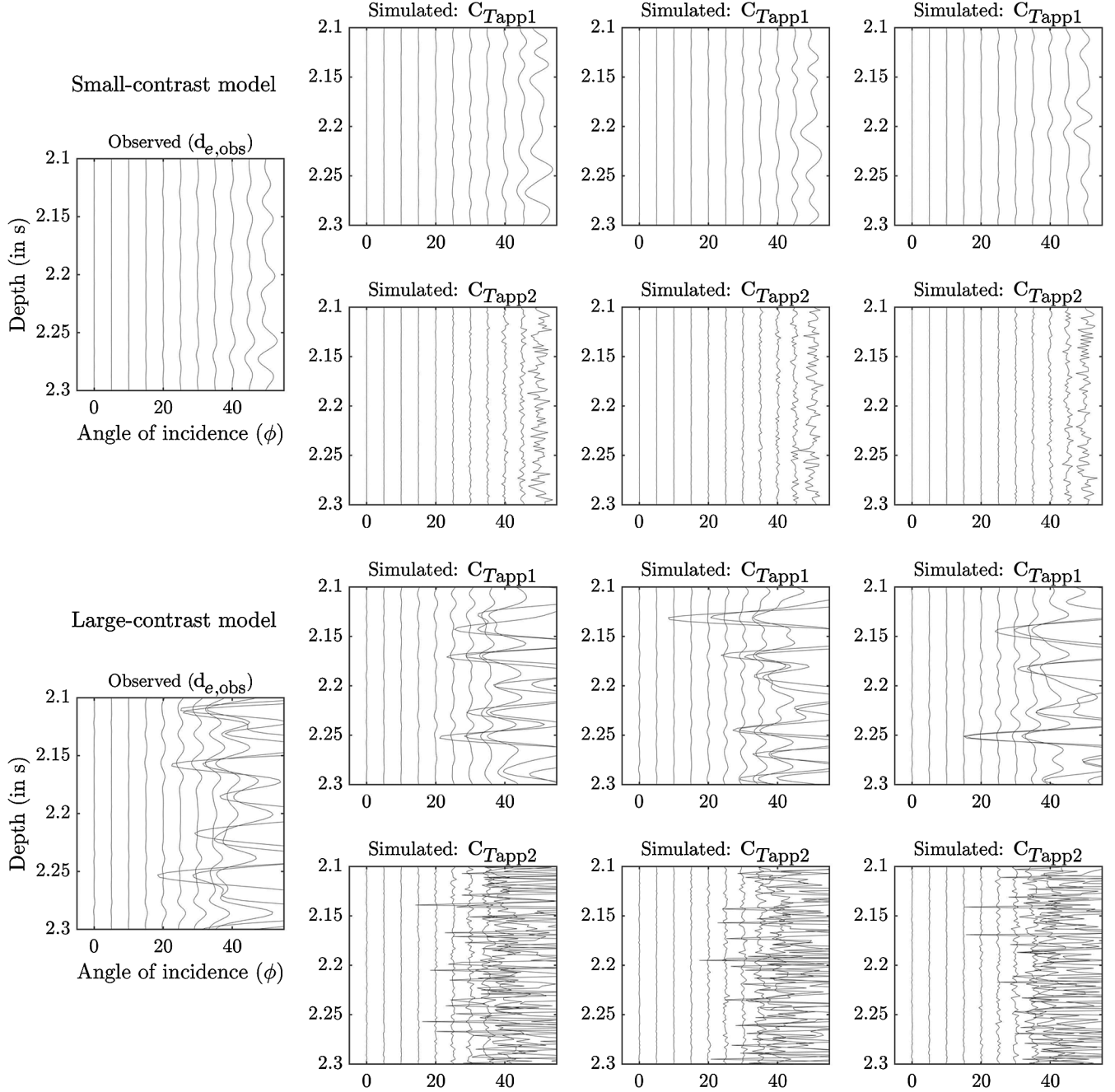


Figure 5. Observed modeling error (left column) plotted against three realizations from a Gaussian random field described by the estimated covariance matrices \mathbf{C}_{Tapp1} and \mathbf{C}_{Tapp2} . The scaling is the same for all realizations.

$$\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp}, \quad (13)$$

where \mathbf{C}_d is the covariance matrix describing the measurement errors. In our case, we let the measurement error be represented by uncorrelated noise; i.e., \mathbf{C}_d is the variance along the diagonal. As demonstrated previously, the linear Buland and Omre forward model is improved using the average angle as opposed to the incidence angle because the forward model is implicitly based on the average angle. AVO angle stacks can, in principle, be processed to reflect the average and incidence angles. But, as the offset-to-angle conversion is based on a typically smooth velocity field (in the present case, a constant field), a typical AVO angle gather will represent an approximation to either the actual incidence or the average angle. Therefore, we will in the remainder of the paper in reference to the linearized inverse problem, refer to an reflection angle. Through a numerical example, the effect of accounting for and discarding the modeling error will now be considered.

Bayesian linearized AVO inversion — Numerical example

The prior realizations in Figure 2 are used as two reference models \mathbf{m}_{ref} . A set of synthetic data is calculated using the Zoeppritz equations for 11 reflection angles (identical to Figure 1). Uncorrelated noise $\varepsilon \sim \mathcal{N}(0, \mathbf{C}_D)$ with $S/N = 5$ between the standard deviation of the forward response and the standard deviation of the noise is added to the data to obtain the “observed” data \mathbf{d}_{obs} .

The small-contrast model presents an ideal case for linearized AVO inversion because the variations in the elastic parameters are smooth (Gaussian), and it is known to be a realization of a Gaussian probability density with known mean and covariance, $\mathcal{N}(\boldsymbol{\mu}_M, \mathbf{C}_M)$. The a priori model is thus known. Because the exact noise model is also known, $\mathcal{N}(0, \mathbf{C}_D)$, the only unknown factor in the linearized AVO inversion method for the smooth prior is the effect of the modeling error.

The large-contrast model cannot be described fully by a Gaussian model, which prohibits assessing the effect of the forward-modeling error in the linearized inversion directly. However, for comparison, a Gaussian prior model has been inferred that best matches the reference model using traditional semivariogram analysis. An exponential type of Gaussian distribution with 0.5 correlation coefficient between the elastic parameters and a range of 12 ms was found. The estimated variances for the elastic parameters are $\sigma_{v_p}^2 = 0.01$, $\sigma_{v_s}^2 = 0.05$, and $\sigma_\rho^2 = 0.002$, respectively. In a real-world setting, a Gaussian model may also not be the obvious choice to describe the prior model, but the Gaussian prior model assumption is needed to make use of equations 11 and 12. Recently, Grana et al. (2017) propose a method that allows using Gaussian prior models with a non-Gaussian 1D marginal distribution. Sabeti et al. (2017) perform direct sequential simulation to allow using non-Gaussian 1D distributions. These methods may allow a better prior model describing models, such as the large-contrast model.

SMALL-CONTRAST MODEL

Figure 6 displays the results from the Bayesian linearized AVO inversion on \mathbf{d}_{obs} from the small-contrast realization. Initially (the top figure), the modeling error is disregarded during inversion, i.e., $\mathbf{C}_D = \mathbf{C}_d$ in equations 11 and 12. This result in some features that seem to be well-resolved but that lie well beyond the 95% confi-

dence interval and represent modeling errors being fitted as data. This bias happens especially where the contrasts in model parameters are fairly high, for instance, at approximately 2150–2220 ms.

The poorest performance is the posterior distribution of the density ρ , where serious bias effects are recognized for large parts of the posterior mean prediction compared with the reference model. This is a worrisome example of an apparently well-resolved feature, which is actually noise (modeling error) being fitted as data. Because the modeling error is increasing for far reflection angles, which are important for the density estimate, this could explain this bias. The P-wave velocity v_p and the velocity ratio v_p/v_s are a bit better resolved, however still showing the noise being fitted as well-resolved features. The acoustic impedance and S-wave velocity are relatively well-resolved, but the reference model is still not fully captured by the uncertainty bands (confidence intervals). In fact, if the 95% confidence intervals were to be an accurate depiction of the uncertainty, one would expect the reference model to be contained within these uncertainty bands at approximately 95% of the samples, which is not the case. This makes the reference model a highly improbable realization of the posterior distribution obtained. Thus, if not accounted for, the modeling error related to the use of the linear Buland and Omre forward is able to create significant biases in inversion results for $S/N = 5$. The lower plots in Figure 6 show the corresponding inversion results accounting for the modeling error using equation 13 and the inferred covariance matrix \mathbf{C}_{Tapp1} shown in Figure 4. The reference model generally lies within the 95% confidence interval. No artifacts are visibly present in the inversion results, and the modeling error seems to be properly accounted for. This is true even at approximately 2160–2180 ms, where the previous nonaccounting posterior prediction failed to resolve the reference model.

By calculating the root-mean-square deviation (rmsd) between the predicted values and the reference model, we can quantify how well each inversion scheme is predicting the reference model:

$$\text{rmsd} = \sqrt{\frac{\sum_{i=1}^n (\mathbf{m}_i - \tilde{\mathbf{m}}_i)^2}{n}}, \quad (14)$$

where \mathbf{m} is the known reference values, $\tilde{\mathbf{m}}$ is the posterior prediction (mean), and n is the number of elements. Because the realization \mathbf{m}_{ref} comes from a Gaussian distribution and the noise distribution is Gaussian, we can also quantify how likely \mathbf{m}_{ref} is as a realization from the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$ by

$$f(\mathbf{m}_{\text{ref}} | \mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)) \sim \exp(-0.5(\mathbf{m}_{\text{ref}} - \tilde{\mathbf{m}})^\top \tilde{\mathbf{C}}_M^{-1}(\mathbf{m}_{\text{ref}} - \tilde{\mathbf{m}})). \quad (15)$$

Given a large sample of \mathbf{m}_{ref} , $\log(f)$ should follow a Gaussian distribution $\mathcal{N}(-N_m/2, \sqrt{N_m}/2)$ if the degrees of freedom of the model parameters are sufficiently high (Hansen et al., 2016). In our case, the expected distribution is $\mathcal{N}(-150, \sqrt{150})$ because $N_m = 300$. In other words, the $\log(f)$ value should be in the interval of -150 ± 25 to be a likely realization from the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$.

The quantitative measures are summarized in Table 1. The rmsd values and the $\log(f)$ value underline the visual results from Figure 6. By including the modeling error in the inversion, rmsd is reduced as much as 40%–50%. The largest improvement is seen

in the density prediction, which comply with the visual results that the poorest prediction of \mathbf{m}_{ref} is offered for the density. Furthermore, the $\log(f)$ value for the case of $\mathbf{C}_D = \mathbf{C}_d$ reveals that \mathbf{m}_{ref} is an very unlikely realization of the posterior distribution $\mathcal{N}(\tilde{\mathbf{m}}, \tilde{\mathbf{C}}_M)$. In other words, the posterior uncertainty does not capture the reference model. By accounting for the modeling error, \mathbf{m}_{ref} can become a highly likely realization from that posterior

distribution, as shown by the $\log(f)$ close to -150 . By calculating the $\log(f)$ value for multiple realizations of \mathbf{m}_{ref} , it is determined that the result is not dependent on the specific realization used in this inversion (see Appendix C). In summary, the inclusion of the estimated modeling error in the linearized AVO inversion for the smooth prior does offer better predictions and a realistic uncertainty band.

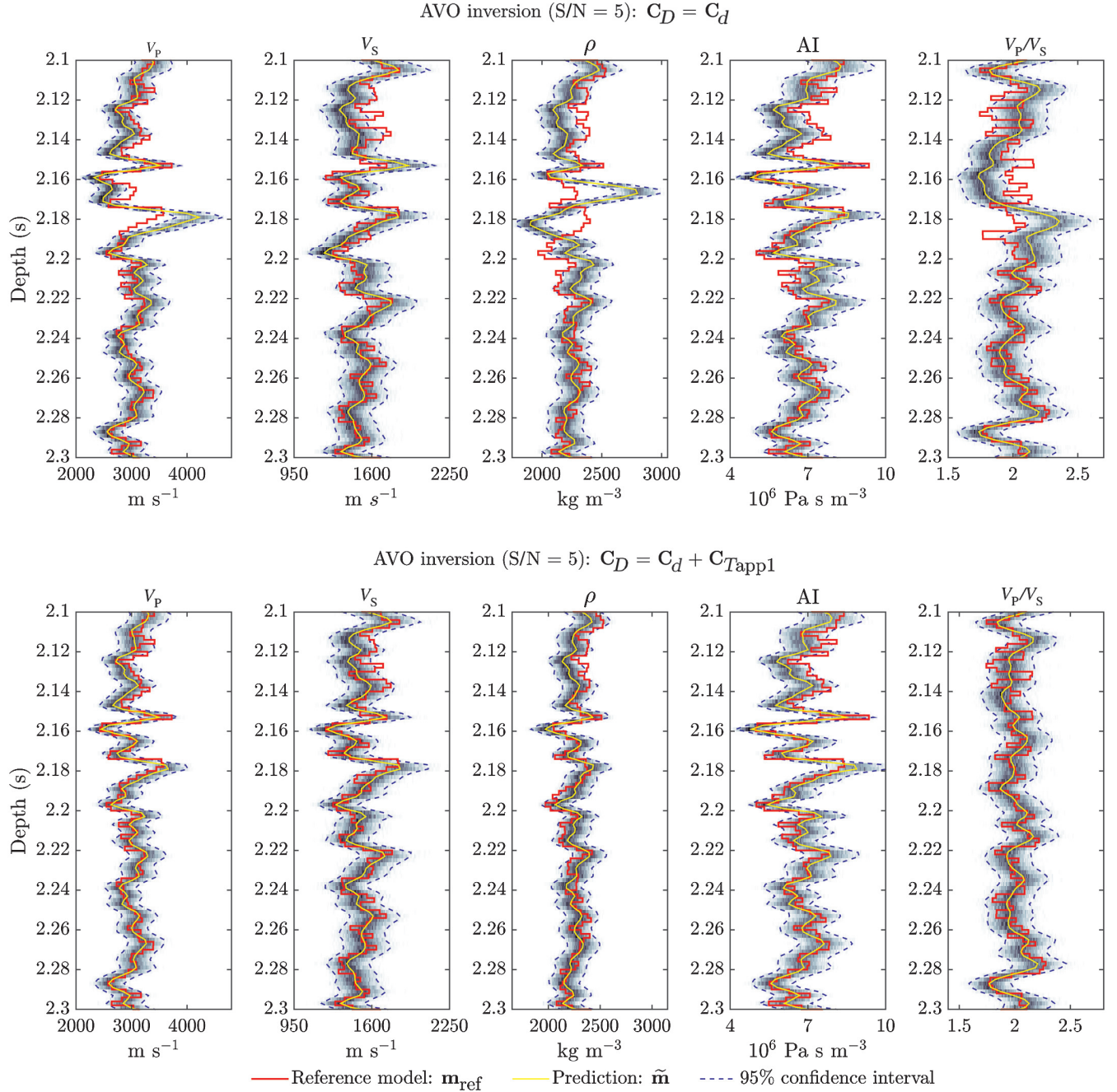


Figure 6. Linear Bayesian inversion performed on reference data calculated from the small-contrast reference model \mathbf{m}_{ref} presented in Figure 2 (red line) with and without accounting for forward modeling errors. Posterior density (black = high density, white = low density), mean (yellow), and 95% confidence interval (dashed blue line) is shown. Before the inversion, the reference data were added random uncorrelated white noise with a standard deviation five times less than the standard deviation of the Zoeppritz forward response of the reference model; i.e., the S/N = 5.

LARGE-CONTRAST MODEL

Figure 7 displays the results from the Bayesian linearized AVO inversion on \mathbf{d}_{obs} from the large-contrast realization and the estimated prior distribution described earlier. For the case of neglecting the forward-modeling error (top figure), i.e., $\mathbf{C}_D = \mathbf{C}_d$, the reference model \mathbf{m}_{ref} is extremely poorly resolved. The reference model is rarely, if ever, within the 95% confidence interval. This happens despite the posterior distribution having a much lower resolution (higher uncertainty) than for the small-contrast prior inversion in Figure 6.

At certain depths, the predicted value is in fact opposite of the reference model. The predictions are for example consistently higher than the reference value for all elastic parameters at approximately 2150 ms. The high-prediction anomaly, which indicates a potential zone of interest, turns out to be an artifact from the inversion. The rmsd values in Table 2 show that predicted values are very far from the reference model for all elastic parameters, especially the S-wave velocity. The mismatch between the predicted values and the reference value is partly expected due to the use of the smooth Gaussian prior model to describe the abrupt changes in elastic properties of the subsurface realization. However, when comparing these results with the case of accounting for the modeling error $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{T_{\text{app1}}}$ (bottom figure), the wrong prior distribution can only explain a certain amount of the misfit. Accounting for the modeling error does, as for the small-contrast case, significantly improve the visual results and the rmsd values. Many of the major trends do seem to be captured by the posterior model, and the uncertainty bands often contain the reference model. Surprisingly, accounting for the modeling error even catches some of the major peaks in the v_P/v_S ratio. The $\log(f)$ values are unsurprisingly low for both inversion cases as seen in Table 2 because we are fitting a discrete reference model with a smooth model. However, the large drop in $\log(f)$ when accounting for the modeling error further emphasizes the above-mentioned points.

DISCUSSION

Many types of inversion of AVO data has been proposed (Simmons and Backus, 1996; Buland and Omre, 2003; Vecken and Da Silva, 2004; Downton, 2005; Rabben et al., 2008; Wilson, 2010; Alemie and Sacchi, 2011; Aune et al., 2013; Grana, 2016). Usually, the data uncertainty is assumed to be Gaussian with zero mean and

a known variance (Rabben et al., 2008; Singleton, 2009; Alemie and Sacchi, 2011; Aune et al., 2013; Grana, 2016). Sometimes, the modeling errors due to the choice of forward model are also considered as part of the general data uncertainty (e.g., Downton, 2005).

It has been proposed to account for systematic errors in Bayesian inversion related to “theory errors” (Riedel et al., 2003; Chen et al., 2007). This is achieved by adding a correlated covariance model describing the modeling uncertainty to the general uncorrelated data uncertainty. This covariance model is obtained using a fixed maximum likelihood estimate (Gerstoft and Mecklenbräuker, 1998). The shape of the correlated covariance modeling errors is in this case assumed to be proportional to the apparent covariance of the data. Alternatively, the data uncertainty including modeling errors has been added as an unknown parameter in a Bayesian inversion framework (Dosso and Holland, 2006).

The presented method of quantifying the forward-modeling error presents a straightforward workflow. In a Bayesian framework, a choice of prior must be performed in all circumstances. A sufficiently large sample of prior realizations and the subsequent modeling error can then be simulated to infer a Gaussian model describing the error. This Gaussian model can be added to the observational uncertainties, and the modeling errors are accounted for as long as the likelihood is Gaussian. Although the investigated errors are not strictly Gaussian, the inversion results indicate that the method offers a significant improvement compared with neglecting the error.

The key benefit of the proposed method is therefore that it allows a correct quantification of uncertainty while avoiding over-fitting the data, i.e., avoiding fitting noise. In addition, the proposed method avoids making any assumptions about the shape of the modeling error, other than it needs to be described by a Gaussian model. In comparison with the previously considered approaches by Riedel et al. (2003) and Chen et al. (2007), this method of quantifying the modeling error is independent of data. The quantification of the modeling error and thus the total data uncertainty can also be described prior to inversion as opposed to inverting for the magnitude of a correlated Gaussian model from data, as proposed by Dosso and Holland (2006).

Handling significant forward-modeling errors for large reflection angle in Bayesian linearized inversion

The results indicate that the forward-modeling error of applying the Buland and Omre forward is significant for larger reflection angle, especially if the prior model has a large contrast in the elastic parameters. This is not surprising because the Buland and Omre forward model depends on a small contrast approximation. The modeling error poses a real concern for reflection angle $\phi > 20^\circ$ if a non-Gaussian prior model is used. For the better case of a Gaussian prior model, our results are in agreement with earlier results suggesting the modeling error is negligible for data sets with $\phi < 30^\circ$ (e.g., Shuey, 1985; Buland and Omre, 2003). Modern collection of seismic data allow processed AVO data to have a reflection angle well exceeding 30° . It is not uncommon that inversion is performed for data sets containing reflection angle of up to 50° (e.g., Barclay et al., 2008).

In practice, when dealing with AVO data with a large reflection angle, (at least) three possible approaches exist.

Table 1. Small-contrast prior inversion. The rmsd calculated using equation 14 for the three elastic parameters v_P , v_S , and ρ divided by the average value to obtain the relative parameters v_P/\bar{v}_P , v_S/\bar{v}_S , $\rho/\bar{\rho}$, and the $\log(f)$ value calculated using equation 15. The values are calculated for both inversion cases ($\mathbf{C}_D = \mathbf{C}_d$ and $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{T_{\text{app1}}}$) shown in Figure 6.

	rmsd	rmsd	rmsd	
\mathbf{C}_D	v_P/\bar{v}_P	v_S/\bar{v}_S	$\rho/\bar{\rho}$	$\log(f)$
\mathbf{C}_d	0.0756	0.0760	0.0775	-1428.7
$\mathbf{C}_d + \mathbf{C}_{T_{\text{app1}}}$	0.0444	0.0573	0.0377	-151.3

Solution 1: Neglect modeling error

One can neglect the possible modeling error of applying an imperfect forward and perform the inversion nonetheless. However, as shown, this strategy leads to significant biases in the posterior distribution compared with the true model. Most concerning is that apparently well-resolved posterior features can in fact, as demonstrated, be due to fitting modeling noise as though it was data. This

can have a fatal effect on subsequent decision making. When random noise on the data is sufficiently large (small S/N), the theoretical errors down in the random noise and the posterior distribution will not be biased significantly from the modeling error. Depending on the subsurface variability, the results in Appendix C suggest that a maximum of $S/N = 0.5$ is used, even for a smooth prior model when the Buland and Omre forward is used in Bayesian linearized AVO inversion for the considered model and configuration.

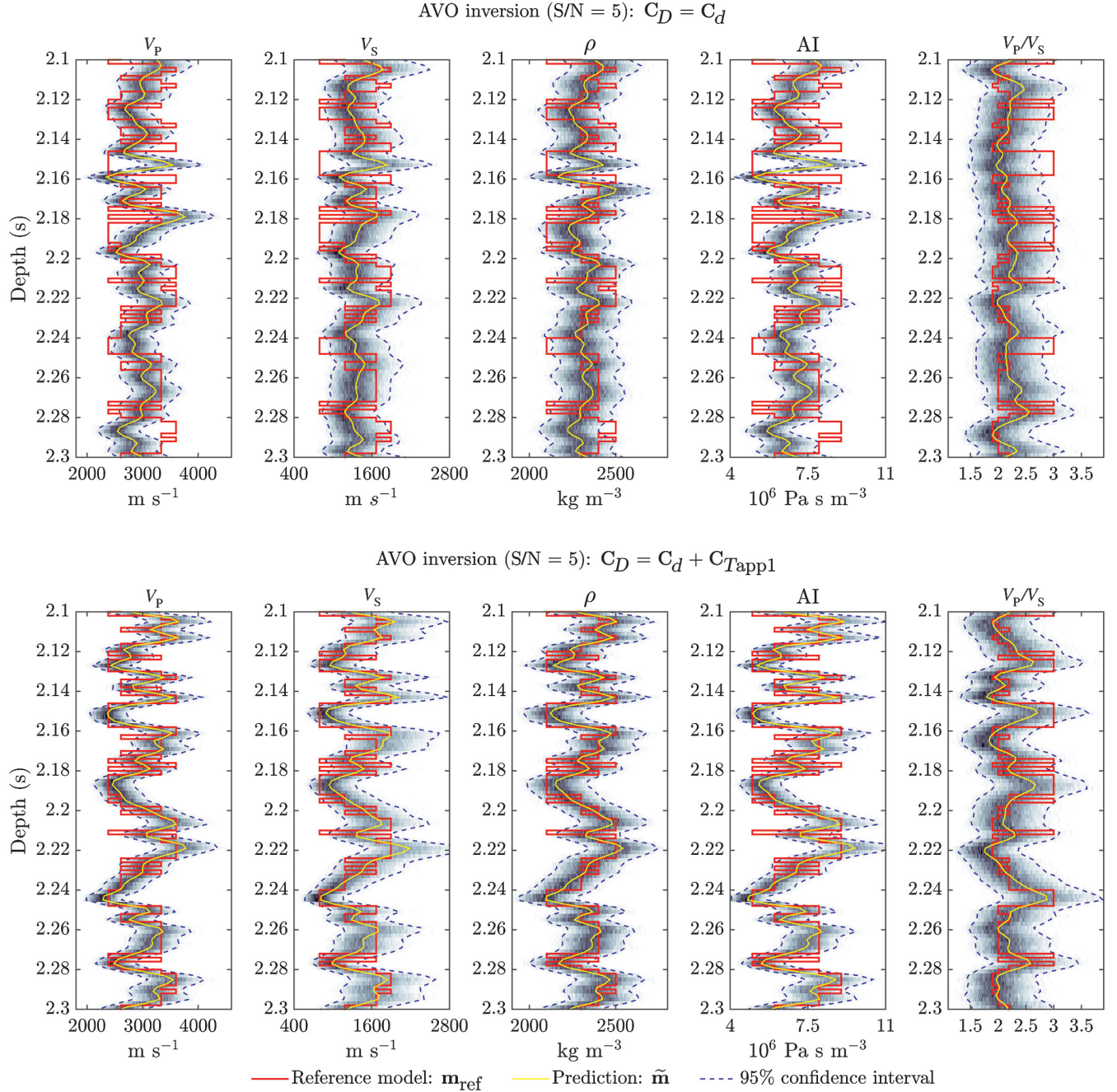


Figure 7. Linear Bayesian inversion performed on reference data calculated from the large-contrast reference model \mathbf{m}_{ref} presented in Figure 2 (red line) with and without accounting for forward modeling errors. Posterior density (black = high density, white = low density), mean (yellow), and 95% confidence interval (dashed blue line) is shown. Before the inversion, the reference data were added random uncorrelated white noise with a standard deviation five times less than the standard deviation of the Zoeppritz forward response of the reference model; i.e., the $S/N = 5$.

Solution 2: Remove larger reflection angle

To avoid introducing significant forward-modeling errors, data from larger reflection angle are neglected (Stolt and Weglein, 1985). Two major issues arise using this strategy. First, seismic data, which are extremely difficult and expensive to acquire, are thrown away. Second, small reflection angles are mostly sensitive to changes in acoustic impedances, whereas larger reflection angles are potentially sensitive to changes in the S- and P-wave velocities (Shuey, 1985). By throwing away the wide reflection angle, significant information about the v_S and v_P/v_S ratio is lost.

Solution 3: Quantify and account for modeling error

Because the effect of the modeling error is significant for $S/N > 0.5$, good data with low noise are therefore explicitly an issue; i.e., the better the data, the bigger the biases in the posterior results. It should once again be stressed that these biases occur even in a best-case scenario with a small-contrast smooth prior. Because the magnitude of the modeling error is higher for the large-contrast prior, one can expect results with even more significant bias. By inferring a Gaussian model directly from a sample of the modeling error as demonstrated here, it has been shown that S/N less than 20 produce reasonable results. If the correlated part is ignored, a $S/N < 2$ has been shown to produce reasonable results (Appendix C). However, because the modeling error is correlated to a large degree, it is rather “naive” to only consider it to be uncorrelated. Our results suggest that the correlated part of the modeling contains significant information, which improves the inversion result.

Future work

In this paper, the focus has been on the modeling error related to the use of a linear approximation of Zoeppritz equations, especially the Buland and Omre forward. This is, as mentioned earlier, just one out of many potential uncertainties regarding AVO data. In the presented Bayesian linear inversion scheme, this modeling error is implicit in the scheme and hence unavoidable. It therefore represents a minimum level of modeling error, which is often neglected. The modeling error associated to for example ignoring anisotropy, imperfect NMO correction, or the use of uncertain wavelets (but to name a few significant sources of modeling errors) may very well be larger than the modeling error associated with the use of a linear small-contrast approximation to Zoeppritz equations, as argued by, e.g., Downton (2005). However, the examples demonstrated that modeling errors of these types can only be ignored when the S/N is

Table 2. Large-contrast prior inversion. The rmsd calculated using equation 14 for the three elastic parameters v_P , v_S , and ρ divided by the average value to obtain the relative parameters v_P/\bar{v}_P , v_S/\bar{v}_S , $\rho/\bar{\rho}$, and the $\log(f)$ value calculated using equation 15. The values are calculated for both inversion cases ($C_D = C_d$ and $C_D = C_d + C_{Tapp1}$) shown in Figure 7.

	rmsd	rmsd	rmsd	
C_D	v_P/\bar{v}_P	v_S/\bar{v}_S	$\rho/\bar{\rho}$	$\log(f)$
C_d	0.1867	0.3229	0.0636	−12893.0
$C_d + C_{Tapp1}$	0.1044	0.2008	0.0382	−282.9

less than 0.5, which is rarely the case in practice of marine seismic data.

The presented methodology can be extended to investigate, quantify, and possibly account for the effect of some of the other AVO modeling errors. Therefore, the next logical step would be to investigate and quantify the effect of the aforementioned sources of AVO modeling error. In particular, the modeling error associated with using the convolutional model as supposed to a NMO-corrected shot gather from waveform simulated seismics. The main requirement is that one must be able to provide a statistical description of the source of the modeling error, from which realizations of the modeling error can be computed. In the present case, the modeling error was linked to the subsurface variability. Further, the methodology can be used to quantify and account for modeling errors also in a full nonlinear, non-Gaussian inversion performed using, e.g., the extended metropolis sampler (see, e.g., Zunino et al., 2015).

A probabilistic Monte Carlo-based sampling strategy would for instance be able to sample the joint posterior distribution for the case of making use of the large-contrast prior (Mosegaard and Tarantola, 1995).

CONCLUSION

We have simulated and quantified the modeling error related to using a linear approximate solution to Zoeppritz equations. The modeling error depends on the degree of subsurface variability and increases with angle of incidence. The Aki and Richards forward is shown to be less inaccurate than the Buland and Omre small-contrast forward, especially using the average angle as opposed to the incidence angle.

A Gaussian model describing this modeling error is inferred from an obtained sample of the modeling error. Realizations from this distribution resemble and mimic the observed values of the modeling error to a degree, where visual discrimination is difficult. A quantitative analysis reveals that the Gaussian model, while not perfect, provides a good description of the generated sample of the modeling error.

A Gaussian model of the modeling error is trivially accounted for as part of linearized AVO inversion. It has been shown that even small modeling errors from a linear forward model, related to a smooth Gaussian prior, can contribute to significant biases in inversion results. Apparently, well-resolved features in the posterior distributions may be caused by fitting modeling errors. In the case there is little subsurface variability expected (assuming a very smooth Gaussian-type a priori covariance model), our results indicate that the modeling error can be disregarded for the considered setup when the S/N of the AVO data is equal to or less than 0.5. Accounting for the modeling error, through the use of the inferred Gaussian model of modeling errors, improves the inversion results dramatically.

If more subsurface variability is expected, the effect of the modeling errors in inversion is even more severe. Our results indicate that accounting for the modeling error in this case provides a more reasonable prediction of the subsurface properties. Furthermore, it limits the amount of major artifacts in the inversion.

ACKNOWLEDGMENTS

We would like to thank H. J. Hansen at Qeye Labs for providing us with realistic properties and real world problems of discrete tran-

sitions in the North Sea. J. Downton for insightful comments during the review process. Calculations were performed using the MATLAB toolbox SIPPI (Hansen et al., 2013). This Ph.D. research is funded by the Innovation Fund Denmark (grant no. 53-2014-3).

APPENDIX A

ESTIMATING MODELING ERROR WITH A GAUSSIAN MODEL

Consider a sample consisting of N realizations of the modeling error ($\mathbf{d}_e^1, \mathbf{d}_e^2, \dots, \mathbf{d}_e^N$). This sample is set up in matrix form as

$$\mathbf{D}_e = [\mathbf{d}_e^1, \mathbf{d}_e^2, \dots, \mathbf{d}_e^N]. \quad (\text{A-1})$$

The mean-modeling error is estimated for each individual point j as

$$d_{T_{\text{app}}}^j = \frac{1}{N} \sum_{i=1}^N \mathbf{D}_e^{i,j}, \quad (\text{A-2})$$

where the mean for the j th data point is the arithmetic mean of all $i = 1, \dots, N$ realizations from the sample. The mean vector for all data points is then achieved by combining the mean of the individual data points calculated in equation A-2:

$$\mathbf{d}_{T_{\text{app}}} = [d_{T_{\text{app}}}^1, d_{T_{\text{app}}}^2, \dots, d_{T_{\text{app}}}^N]. \quad (\text{A-3})$$

The covariance of the modeling error is estimated as

$$\mathbf{C}_{T_{\text{app}}} = \frac{1}{N} [\mathbf{D}_e - \mathbf{D}_{T_{\text{app}}}][\mathbf{D}_e - \mathbf{D}_{T_{\text{app}}}]^T, \quad (\text{A-4})$$

where $\mathbf{D}_{T_{\text{app}}} = [\mathbf{d}_{T_{\text{app}}}^T, \mathbf{d}_{T_{\text{app}}}^T, \dots, \mathbf{d}_{T_{\text{app}}}^T]$ is a matrix containing N repetitions of the mean vector of equation A-3. The mean $\mathbf{d}_{T_{\text{app}}}$ does in practice tend toward zero for large N as the modeling error is either negative or positive depending on the elastic parameters, which have a fixed mean. The $\mathbf{D}_{T_{\text{app}}}$ can therefore be excluded from equation A-4 for large N .

APPENDIX B

QUANTITATIVE ASSESSMENT OF GAUSSIAN MODEL

A quantitative measure of the validity of the Gaussian assumption on the modeling error can be obtained by analyzing how probable the simulated modeling errors are as a realization of the inferred Gaussian model. The definition of the Gaussian model on a modeling error realization \mathbf{d}_e is

$$f(\mathbf{d}_e | \mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})) \sim \exp(-0.5(\mathbf{d}_e - \mathbf{d}_{T_{\text{app}}})^T \mathbf{C}_{T_{\text{app}}}^{-1} (\mathbf{d}_e - \mathbf{d}_{T_{\text{app}}}), \quad (\text{B-1})$$

$$= \exp(-0.5 \Omega), \quad (\text{B-2})$$

where $\Omega = (\mathbf{d}_e - \mathbf{d}_{T_{\text{app}}})^T \mathbf{C}_{T_{\text{app}}}^{-1} (\mathbf{d}_e - \mathbf{d}_{T_{\text{app}}})$ will then be distributed according to a χ^2 -distribution with N_d number of degrees of freedom, which for large N_d tend to follow a Gaussian distribution $\Omega \sim \mathcal{N}(N_d, 2N_d)$ (Tarantola, 2005). This also means that $\log(f(\mathbf{d}_e | \mathcal{N}(\mathbf{d}_{T_{\text{app}}}, \mathbf{C}_{T_{\text{app}}})) = -0.5\Omega$, for large N_d , will tend to follow a Gaussian distribution $\mathcal{N}(-N_d/2, \sqrt{N_d}/2)$ (Hansen et al., 2016). Specifically, in our example, $\log(f)$ should follow the Gaussian distribution $\mathcal{N}(-550, \sqrt{550})$ because $N_d = 1100$.

Figure B-1 shows the histograms of $\log(f(\mathbf{d}_e | \mathcal{N}(\mathbf{d}_{T_{\text{app}}1}, \mathbf{C}_{T_{\text{app}}1})))$ for actual simulated modeling errors ($\mathbf{d}_{e,\text{obs}}$), realizations of $\mathcal{N}(\mathbf{d}_{T_{\text{app}}1}, \mathbf{C}_{T_{\text{app}}1})$ ($\mathbf{d}_{e,\text{app}1}$) and samples of $\mathcal{N}(\mathbf{d}_{T_{\text{app}}2}, \mathbf{C}_{T_{\text{app}}2})$ ($\mathbf{d}_{e,\text{app}2}$). The histogram of $\log(f(\mathbf{d}_{e,\text{app}1}))$ values (dark gray) follows the Gaussian distribution as described above, with a mean approximately -550 . This is expected as we are comparing realizations from $\mathbf{C}_{T_{\text{app}}1}$ with itself.

If the observed modeling error is described perfectly by the estimated Gaussian covariance model, the sample of the observed modeling error (black) would plot on top of the $\mathbf{C}_{T_{\text{app}}1}$ sample (dark gray)

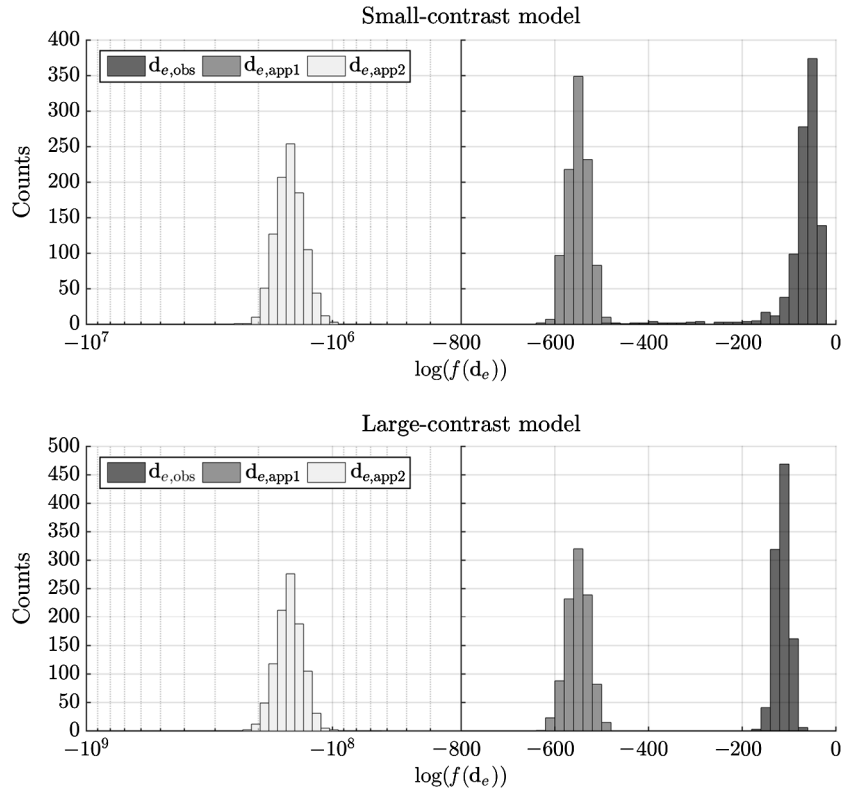


Figure B-1. Log-likelihood values calculated for 1000 realizations of the observed modeling error (black) and the two samples of $\mathbf{C}_{T_{\text{app}}1}$ (dark gray), $\mathbf{C}_{T_{\text{app}}2}$ (light gray), respectively. The results are shown for the two types of prior models, (top row) the small-contrast prior distribution and (base row) the large-contrast prior distribution. Notice the logarithmic scale used on the left side plots.

in Figure B-1. The histogram of the observed modeling error for the small-contrast prior is left-skewed with values mostly lying at approximately -60 with a tail in the range between -200 and -500 . In the large-contrast case, a tail is not present, but the distribution is otherwise similar, with values of approximately -120 . Such consistently higher log-likelihood values suggest that the sample values are consistently closer to the mean (zero) value than is expected from a normal distribution. For comparison, the log-likelihood histograms of realizations from the uncorrelated Gaussian model C_{Tapp2} (light gray) show values exceeding the previous two with several orders of magnitude. Therefore, the uncorrelated Gaussian model C_{Tapp2} is not a very good description of the actual noise sample, whereas the correlated Gaussian C_{Tapp1} while not perfect, does a much better job.

Figure B-2 shows the 1D marginal distribution of the observed modeling error $\mathbf{d}_{e,obs}$ (dashed black) compared with the estimated models below (Figure B-2a and B-2d) and greater than $\phi = 30^\circ$ (Figure B-2b and B-2e) for the two subsurface models. The distribution of $\mathbf{d}_{e,app1}$ (dark gray) and $\mathbf{d}_{e,app2}$ (dotted light gray) shows identical distributions on both histograms because they are both from a Gaussian model with a similar variance. Most noticeable for the larger angles of incidence and for the large-contrast prior distribution in general, Figure B-2b, B-2d, and B-2e demonstrates that the observed modeling error produce a slimmer 1D distribution, with longer tails that cannot be completely described by the Gaussian model. This slight discrepancy to the Gaussian model explains the discrepancy observed in Figure B-1.

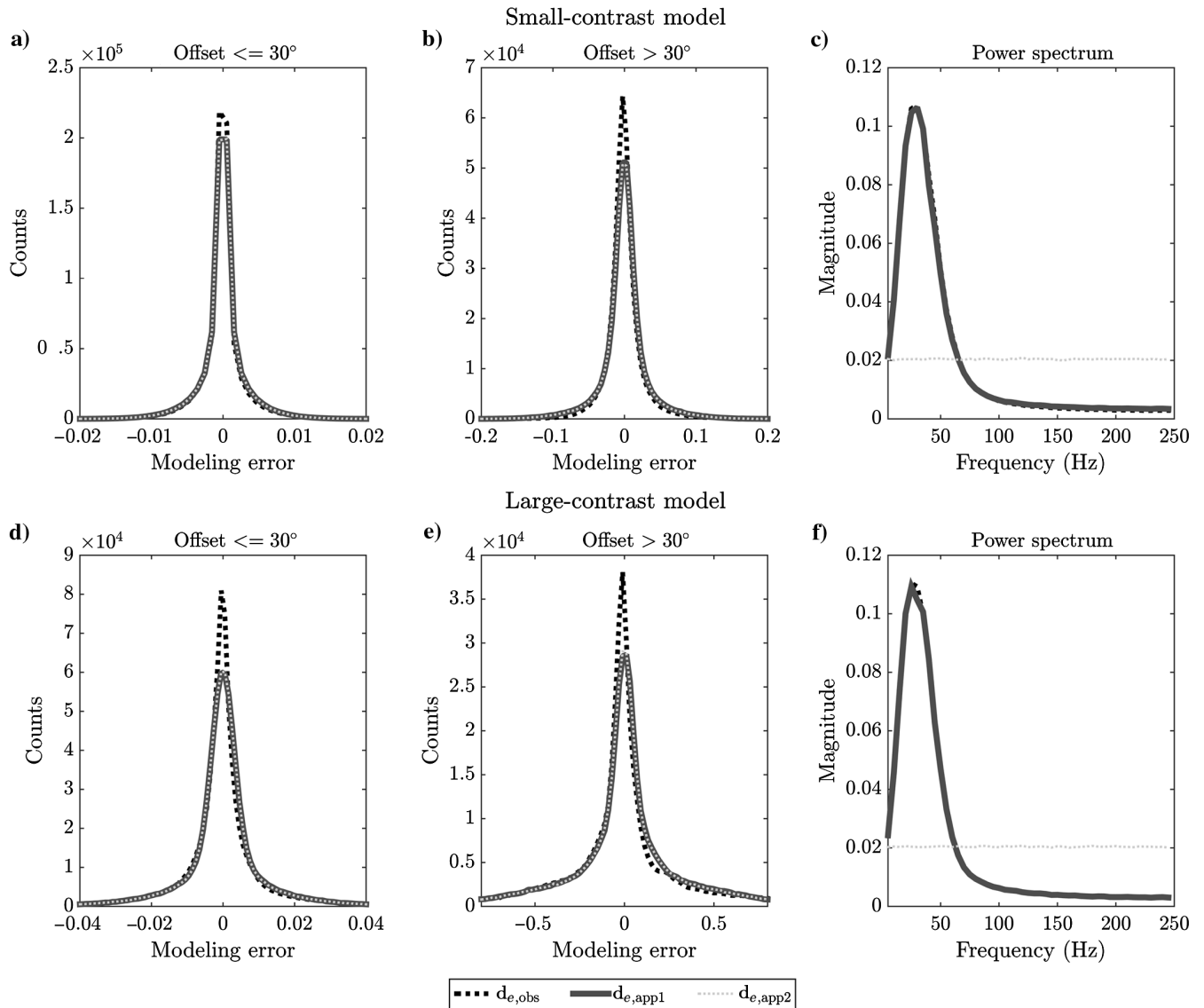


Figure B-2. (a, b, d, and e) Histogram of modeling error and (c and f) power spectrum shown for three samples consisting of 1000 realizations of the observed modeling errors $\mathbf{d}_{e,obs}$ (dashed black), $\mathbf{d}_{e,app1}$ (dark gray), and $\mathbf{d}_{e,app2}$ (light dots), respectively. The histograms and power spectrum are shown for (a-c) the small-contrast prior distribution and (d-f) the large-contrast prior distribution. The histograms of the modeling error are split between angles of incidence above and below $\phi = 30^\circ$. The final column shows a normalized frequency spectrum of all angle of incidence traces.

In addition, a power spectral density (power spectrum) is calculated using a fast Fourier transformation on the three samples. The normalized power spectrum for the average of all incidence angles is shown in Figure B-2c and B-2f. Realizations from the estimated Gaussian model $\mathbf{d}_{e,app1}$ show a similar frequency pattern to the observed modeling error $\mathbf{d}_{e,obs}$, in which frequencies at approximately 30 Hz contain the most power. On the contrary, all frequencies, as expected, are equally represented for the sample of the uncorrelated Gaussian model $\mathbf{d}_{e,app2}$. This result underpins the qualitative result from Figure 5, where visual discrimination between the observed modeling error and realizations from \mathbf{C}_{Tapp1} is difficult, if not impossible.

APPENDIX C

MULTIPLE INVERSION RESULTS

To determine whether the bias found in Figure 6 was an isolated or a more general problem, we repeat the inversion 1000 times, using 1000 different realizations of the small-contrast prior as reference models. The Zoeppritz forward response for each realization is calculated and noise is added, as described previously, to obtain 1000 observed reference AVO data sets.

If the estimated posterior Gaussian probability density adequately describes the solution to the inverse problem, then each of the 1000 reference models should be realistic realizations of the corresponding posterior Gaussian probability. This is quantified using equation 15 for each realization and the corresponding posterior distribution.

Figure C-1 shows histograms of the distribution of $\log(f)$ for pairs of 1000 reference models, and their corresponding Gaussian posterior distributions. Three different noise models (\mathbf{C}_D) are used for inverting the data. In the first case (left column), the noise model is equal to the observational uncertainty $\mathbf{C}_D = \mathbf{C}_d$ (i.e., modeling errors are ignored). For $S/N = 5$, the histogram is not following the expected Gaussian distribution (black line) and the $\log(f)$ tend to have much lower values. These outliers correspond to reference models that are inconsistent with the posterior distribution. Only a few prior realizations are found within the expected distribution. This confirms the result of heavy biases from the single prior realization in Figure 6. To avoid biases completely, the results from the likelihood histograms indicate that only $S/N = 1$ or less should be used if only the uncorrelated observational uncertainty is used as data uncertainty.

In the second case (the middle column), the data uncertainty consists of the observational uncertainty and the inferred Gaussian model for the modeling error $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp1}$. For $S/N = 5$, the histogram is shifted to more or less fit inside the expected distribution. By accounting for the modeling error, the bias in results has more or less vanished. This confirms the visual results from Figure 6. The distribution of $\log(f)$ in Figure C-1 suggests that by accounting for the modeling error, it is possible to drastically im-

prove inversion results in terms of the $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp1}$ being an accurate quantification of the uncertainty. Even for $S/N = 20$, the inversion results indicate that the proposed data uncertainty is a relatively accurate description of the uncertainty.

Finally, the inversion is also performed with data uncertainty $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp2}$, i.e., the assumption of uncorrelated modeling errors. The resulting histograms (the right column) show a similar distribution as for using $\mathbf{C}_D = \mathbf{C}_d$. In contrast to the first case, the distribution of $\log(f)$ generally attains higher values for $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp2}$. In practice, this means that adding the uncorrelated Gaussian model, the inversion will provide a posterior distribution inconsistent with the reference model, though not to the same extent as for the observational uncertainties only. To avoid biases in the inversion results, it is here indicated that an $S/N = 2$ should not be exceeded. Because \mathbf{C}_{Tapp2} is basically just white noise with differing variance, adding this to data uncertainty is somewhat comparable with just increasing the magnitude of the observational uncertainties. Therefore, the resolution (variance of the posterior) of $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp2}$ is also lower than for the case of $\mathbf{C}_D = \mathbf{C}_d$. This could explain why the results are generally better because the lowered resolution allows for reference models to be a realistic realization from the inversion result.

At $S/N = 0.5$, the observational data noise is large enough such that the distribution of $\log(f)$ is similar for all three uncertainty

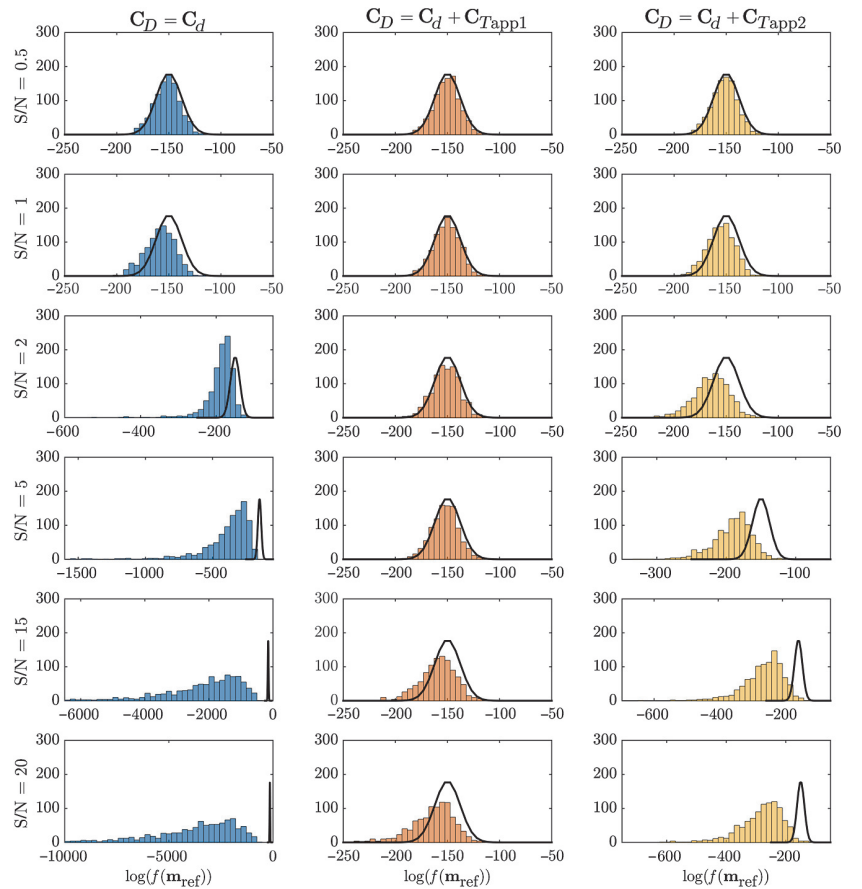


Figure C-1. Histograms of log-likelihood values according to equation 15 for a sample of 1000 prior realizations at different S/N levels. The linear Bayesian AVO inversion has been performed for three data uncertainty models $\mathbf{C}_D = \mathbf{C}_d$, $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp1}$, and $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_{Tapp2}$.

models. The effect of the modeling error is therefore insignificant for $S/N \leq 0.5$. This refers to unrealistic amounts of noise on the data and thus to where the resolution of the posterior models is poor. In other words, this demonstrates that the modeling error related to using the linearized Buland and Omre forward for reflection angle up to 50° can only be safely ignored if the S/N is extremely low.

REFERENCES

- Aki, K., and P. G. Richards, 1980, Quantitative seismology, theory and methods: Vol I and II, 1st ed.: W.H. Freedman Publishing.
- Alemie, W., and M. D. Sacchi, 2011, High-resolution three-term AVO inversion by means of a trivariate Cauchy probability distribution: *Geophysics*, **76**, no. 3, R43–R55, doi: [10.1190/1.3554627](https://doi.org/10.1190/1.3554627).
- Armstrong, M., A. Galli, G. Le Loc'h, D. Renard, B. Deligez, F. Geffroy, and R. Eschard, 2011, Plurigaussian simulations in geosciences: Springer.
- Aune, E., J. Eidsvik, and B. Ursin, 2013, Three-dimensional non-stationary and non-linear isotropic AVO inversion: *Geophysical Journal International*, **194**, 787–803, doi: [10.1093/gji/ggt127](https://doi.org/10.1093/gji/ggt127).
- Ball, V., L. Tenorio, C. Schiott, J. P. Blangy, and M. Thomas, 2015, Uncertainty in inverted elastic properties resulting from uncertainty in the low-frequency model: *The Leading Edge*, **34**, 1028–1035, doi: [10.1190/tle34091028.1](https://doi.org/10.1190/tle34091028.1).
- Barclay, F., A. Bruun, K. B. Rasmussen, J. C. Alfaro, A. Cooke, D. Cooke, D. Salter, R. Godfrey, D. Lowden, S. McHugo, H. Ozdemir, S. Pickering, F. G. Pineda, J. Herwanger, S. Volterrani, A. Murineddu, A. Rasmussen, and R. Roberts, 2008, Seismic inversion: Reading between the lines: *Oil-field Review*, **20**, 42–63.
- Bosch, M., T. Mukerji, and E. F. Gonzalez, 2010, Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review: *Geophysics*, **75**, no. 5, 75A165–75A176, doi: [10.1190/1.3478209](https://doi.org/10.1190/1.3478209).
- Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: *Geophysics*, **68**, 185–198, doi: [10.1190/1.1543206](https://doi.org/10.1190/1.1543206).
- Castagna, J. P., and M. M. Backus, 1993, Offset-dependent reflectivity: Theory and practice of AVO analysis: SEG.
- Castagna, J. P., H. W. Swan, and D. J. Foster, 1998, Framework for AVO gradient and intercept interpretation: *Geophysics*, **63**, 948–956, doi: [10.1190/1.1444406](https://doi.org/10.1190/1.1444406).
- Chen, M. A., M. Riedel, R. D. Hyndman, and S. E. Dosso, 2007, AVO inversions of BSRs in marine gas hydrate studies: *Geophysics*, **72**, no. 2, C31–C43, doi: [10.1190/1.2435604](https://doi.org/10.1190/1.2435604).
- Claerbout, J. F., C. Green, and I. Green, 2004, Earth soundings analysis: Processing versus inversion: Stanford University.
- Cooke, D., and W. A. Schneider, 1983, Generalized linear inversion of reflection seismic data: *Geophysics*, **48**, 665–676, doi: [10.1190/1.1441497](https://doi.org/10.1190/1.1441497).
- Dosso, S. E., and C. W. Holland, 2006, Geoacoustic uncertainties from viscoelastic inversion of seabed reflection data: *IEEE Journal of Oceanic Engineering*, **31**, 657–671, doi: [10.1109/JOE.2005.858358](https://doi.org/10.1109/JOE.2005.858358).
- Downton, J. E., 2005, Seismic parameter estimation from AVO inversion: Ph.D. thesis, University of Calgary.
- Downton, J. E., and C. P. Ursenbach, 2006, Linearized amplitude variation with offset (AVO) inversion with supercritical angles: *Geophysics*, **71**, no. 5, E49–E55, doi: [10.1190/1.2227617](https://doi.org/10.1190/1.2227617).
- Gerstoft, P., and C. F. Mecklenbräuker, 1998, Ocean acoustic inversion with estimation of a posteriori probability distributions: *Journal of Acoustical Society America*, **104**, 808–819, doi: [10.1121/1.423355](https://doi.org/10.1121/1.423355).
- Grana, D., 2016, Bayesian linearized rock-physics inversion: *Geophysics*, **81**, no. 6, D625–D641, doi: [10.1190/geo2016-0161.1](https://doi.org/10.1190/geo2016-0161.1).
- Grana, D., T. Fjeldstad, and H. Omre, 2017, Bayesian Gaussian mixture linear inversion for geophysical inverse problems: *Mathematical Geosciences*, **49**, 1–37, doi: [10.1007/s11004-016-9658-6](https://doi.org/10.1007/s11004-016-9658-6).
- Hansen, T. M., K. S. Cordua, B. H. Jacobsen, and K. Mosegaard, 2014, Accounting for imperfect forward modeling in geophysical inverse problems: Exemplified for crosshole tomography: *Geophysics*, **79**, no. 3, H1–H21, doi: [10.1190/geo2013-0215.1](https://doi.org/10.1190/geo2013-0215.1).
- Hansen, T. M., K. S. Cordua, M. C. Looms, and K. Mosegaard, 2013, SIPPI: A Matlab toolbox for sampling the solution to inverse problems with complex prior information. Part 1: Methodology: *Computers and Geosciences*, **52**, 470–480, doi: [10.1016/j.cageo.2012.09.004](https://doi.org/10.1016/j.cageo.2012.09.004).
- Hansen, T. M., K. S. Cordua, A. Zunino, and K. Mosegaard, 2016, Probabilistic integration of geo-information, in M. Moorkamp, P. G. Lelièvre, N. Linde, and A. Khan, eds., *Integrated imaging of the earth: Theory and applications*, 1st ed.: John Wiley & Sons, 93–116.
- Lay, T., and T. C. Wallace, 1995, *Modern global seismology*, 1st ed.: Academic Press.
- Li, L., J. Caers, and P. Sava, 2015, Assessing seismic uncertainty via geostatistical velocity-model perturbation and image registration: An application to subsalt imaging: *The Leading Edge*, **34**, 1064–1070, doi: [10.1190/tle34091064.1](https://doi.org/10.1190/tle34091064.1).
- Mavko, G., T. Mukerji, and J. Dvorkin, 2009, *The rock physics handbook*, 2nd ed.: Cambridge University Press.
- Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: *Journal of Geophysical Research*, **100**, 12431–12447, doi: [10.1029/94JB03097](https://doi.org/10.1029/94JB03097).
- Mosegaard, K., and A. Tarantola, 2002, 16 Probabilistic approach to inverse problems, in W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, eds., *International handbook of earthquake and engineering seismology*, Part A: Academic Press, 237–265.
- Rabben, T. E., H. Tjelmeland, and B. Ursin, 2008, Non-linear Bayesian joint inversion of seismic reflection coefficients: *Geophysical Journal International*, **173**, 265–280, doi: [10.1111/j.1365-246X.2007.03710.x](https://doi.org/10.1111/j.1365-246X.2007.03710.x).
- Riedel, M., S. E. Dosso, and L. Beran, 2003, Uncertainty estimation for amplitude variation with offset (AVO) inversion: *Geophysics*, **68**, 1485–1496, doi: [10.1190/1.1620621](https://doi.org/10.1190/1.1620621).
- Sabeti, H., A. Moradzadeh, F. D. Ardejani, L. Azevedo, A. Soares, P. Pereira, and R. Nunes, 2017, Geostatistical seismic inversion for non-stationary patterns using direct sequential simulation and co-simulation: *Geophysical Prospecting*, doi: [10.1111/1365-2478.12502](https://doi.org/10.1111/1365-2478.12502).
- Sen, M. K., and P. L. Stoffa, 1996, Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion: *Geophysical Prospecting*, **44**, 313–350, doi: [10.1111/j.1365-2478.1996.tb00152.x](https://doi.org/10.1111/j.1365-2478.1996.tb00152.x).
- Shuey, R., 1985, A simplification of the Zoeppritz equations: *Geophysics*, **50**, 609–614, doi: [10.1190/1.1441936](https://doi.org/10.1190/1.1441936).
- Simmons, J. L., and M. M. Backus, 1996, Waveform-based AVO inversion and AVO prediction-error: *Geophysics*, **61**, 1575–1588, doi: [10.1190/1.1444077](https://doi.org/10.1190/1.1444077).
- Singleton, S., 2009, The effects of seismic data conditioning on prestack simultaneous impedance inversion: *The Leading Edge*, **28**, 772–781, doi: [10.1190/1.3167776](https://doi.org/10.1190/1.3167776).
- Stolt, R. H., and A. B. Weglein, 1985, Migration and inversion of seismic data: *Geophysics*, **50**, 2458–2472, doi: [10.1190/1.1441877](https://doi.org/10.1190/1.1441877).
- Svendsen, J., E. Esmerode, H. Hansen, M. Lykker, H. Friis, and T. Størmosse, 2012, Unravelling the injected nature of a paleocene sand system: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, 4.
- Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation, 1st ed.: SIAM.
- Tarantola, A., and B. Valette, 1982a, Generalized nonlinear inverse problems solved using the least squares criterion $X = pt$: *Reviews of Geophysics and Space Physics*, **20**, 219–232, doi: [10.1029/RG020i002p00219](https://doi.org/10.1029/RG020i002p00219).
- Tarantola, A., and B. Valette, 1982b, Inverse problems = quest for information: *Journal of Geophysics*, **50**, 159–170.
- Thore, P., 2015, Uncertainty in seismic inversion: What really matters?: *The Leading Edge*, **34**, 1000–1004, doi: [10.1190/tle34091000.1](https://doi.org/10.1190/tle34091000.1).
- Vecken, P. C. H., and M. Da Silva, 2004, Seismic inversion methods and some of their constraints: *First Break*, **22**, 47–70, doi: [10.3997/1365-2397.2004011](https://doi.org/10.3997/1365-2397.2004011).
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Wilson, A., 2010, Theory and methods of frequency-dependent AVO inversion: Ph.D. thesis, University of Edinburgh.
- Zoeppritz, K., 1919, Über Reflexion und Durchgang seismischer Wellen durch Unstetigkeitsflächen: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, **1919**, 66–84.
- Zunino, A., K. Mosegaard, K. Lange, Y. Melnikova, and T. M. Hansen, 2015, Monte Carlo reservoir analysis combining seismic reflection data and informed priors: *Geophysics*, **80**, no. 1, R31–R41, doi: [10.1190/geo2014-0052.1](https://doi.org/10.1190/geo2014-0052.1).